

Challenges in S(P)DEs and Bayesian inference

Simon J.A. Malham and Anke Wiese
(Heriot–Watt University, Edinburgh)

Imperial 2009: May 11–13th

Brown report

Applied mathematics at the US Department of Energy:

Develop new approaches for efficient modeling of large stochastic systems.

... particularly spatially dependent systems.

Invest in analysis and algorithms for stochastic optimization.

How do we modify models of the atmosphere and clouds to incorporate newly collected data of possibly of new types?

Stochastic differential equations

$$dy_t = \tilde{V}_0(y_t) dt + V_1(y_t) dW_t^1 + \cdots + V_d(y_t) dW_t^d$$

Four approaches to approximation, solve:

- ▶ PDE (transition probability distribution)
- ▶ for a weak approximation (Monte–Carlo)
- ▶ for a strong approximation (Monte–Carlo)
- ▶ pathwise (rough paths)

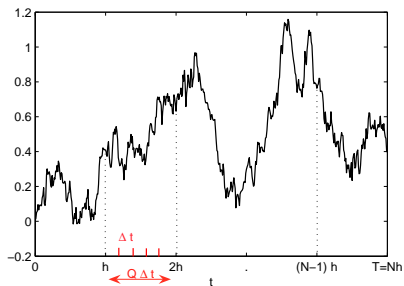
First three: expectation and higher moments of solution sought.

Basic setting

$$y_t = y_0 + \int_0^t \tilde{V}_0(y_\tau) d\tau + \sum_{i=1}^d \int_0^t V_i(y_\tau) dW_\tau^i$$

- ▶ Itô form
- ▶ Non-commuting vector fields: $V_i = \sum_{j=1}^n V_i^j(y) \partial_{y_j}$
- ▶ d -dimensional driving signal: (W^1, \dots, W^d)
- ▶ Convention: $W_t^0 \equiv t$
- ▶ Solution process: $y: \mathbb{R}_+ \rightarrow \mathbb{R}^N$

Wiener process



- ▶ $W_t - W_s \sim N(0, \sqrt{t - s})$
- ▶ Independent increments
- ▶ Continuous, nowhere differentiable

Applications

- ▶ *Finance*: Heston model for pricing options. Stock price is a stochastic process u with stochastic volatility v :

$$\begin{aligned} du_t &= \mu u_t dt + \sqrt{v_t} u_t dW_t \\ dv_t &= \kappa(\theta - v_t) dt + \varepsilon \sqrt{v_t} dZ_t \end{aligned}$$

- ▶ Molecular simulation: Langevin dynamics
- ▶ DNA damage dynamics (Chickarmane *et al.*)
- ▶ Neuronal dynamics (Coombes & Lord)
- ▶ Biochemical reactions (Burrage)
- ▶ Ocean/weather modelling: Bayesian inference (Stuart, Jones)
- ▶ Oil extraction: porous media
- ▶ Subspace tracking: inference on manifolds (Srivastava)

Itô's lemma

$$dy = \tilde{V}_0 \circ y dt + \sum_{i=1}^d V_i \circ y dW^i$$

$$\begin{aligned} \Rightarrow d(f \circ y) &= \tilde{V}_0 \cdot \partial_y(f \circ y) dt + \sum_{i=1}^d V_i \cdot \partial_y(f \circ y) dW^i \\ &\quad + \frac{1}{2} \sum_{i=1}^d (V_i \otimes V_i) : \partial_{yy}(f \circ y) \text{ "dt" } \end{aligned}$$

$$\Rightarrow d(f \circ y) = \mathcal{L}(f \circ y) dt + \sum_{i=1}^d V_i \cdot \partial_y(f \circ y) dW_t^i$$

where

$$\mathcal{L} := \tilde{V}_0 \cdot \partial_y + \frac{1}{2} \sum_{i=1}^d (V_i \otimes V_i) : \partial_{yy}$$

Related PDE

- ▶ Consider Itô SDE for $f \circ y_t$ with initial data $y_0 \equiv y$:

$$\begin{aligned} \Rightarrow f \circ y_t &= f \circ y + \int_0^t \mathcal{L}(f \circ y_\tau) d\tau \\ &\quad + \sum_{i=1}^d \int_0^t V_i \cdot \partial_y(f \circ y_\tau) dW_\tau^i \end{aligned}$$

- ▶ Feynman–Kac \Rightarrow solution of PDE

$$\partial_t u = \mathcal{L} u \quad \text{with} \quad u(0, y) = f(y)$$

$$\text{is } u(t, y) = \mathbb{E}(f(y_t))$$

Black–Scholes–Merton PDE

- ▶ Recall:

$$\mathcal{L} := \tilde{V}_0 \cdot \partial_y + \frac{1}{2} \sum_{i=1}^d (V_i \otimes V_i) : \partial_{yy}$$

- ▶ Constant volatility v , stock/index value u_t evolves:

$$du_t = \mu u_t dt + \sqrt{v} u_t dW_t$$

- ▶ Current price of option at time t is $C(t) = f(t, u_t)$
- ▶ Itô formula and financial argument to duplicate C by a portfolio consisting of investment of u and a bond (risk-free with interest rate r) \Rightarrow

$$\partial_t f + ru \partial_u f + \frac{1}{2} v u^2 \partial_{uu} f - r f = 0$$

Weak approximation

- ▶ Replace Gaussian increments $\Delta W^i(t_n, t_{n+1})$ by simpler RVs $\Delta \hat{W}^i(t_n, t_{n+1})$ with appropriate moment properties, eg. by branching process:

$$P(\Delta \hat{W}^i(t_n, t_{n+1}) = \pm\sqrt{h}) = \frac{1}{2}$$

- ▶ Expectation of approximation \hat{y}_T across all paths at the final time T is close to the expectation of the true solution:

$$\|\mathbb{E}(y_T) - \mathbb{E}(\hat{y}_T)\| = \mathcal{O}(h^p)$$

- ▶ No pathwise comparison: paths not “close” to Wiener paths

Strong approximation

- ▶ Generate approximate Wiener process paths
- ▶ Pick independent increments from

$$\Delta W^i(t_n, t_{n+1}) \sim N(0, \sqrt{h})$$

- ▶ Since we have followed Wiener path approximations, we can compare y_T with \hat{y}_T , they're close in the sense:

$$\mathbb{E} \|y_T - \hat{y}_T\| = \mathcal{O}(h^{\frac{p}{2}})$$

Rough paths (Lyons)

- ▶ Does solution depend continuously on driving path?
- ▶ Lévy area $A_{ij} \equiv \frac{1}{2}(J_{ij} - J_{ji})$
- ▶ V_i non-commuting $\Rightarrow y_t$ depends continuously on W_i, A_{ij}
- ▶ i.e. $(\hat{W}_i, \hat{A}_{ij}) \rightarrow (W_i, A_{ij}) \Rightarrow \hat{y} \rightarrow y$
- ▶ A_{ij} not included \exists example: $W_i \rightarrow 0$ but $\hat{y} \not\rightarrow 0$
- ▶ Back to strong solutions:

$$\hat{W}^i \rightarrow W^i \quad \Rightarrow \quad \mathbb{E} \|\hat{y}_t - y_t\|_{L^2}^2 \rightarrow 0$$

Stochastic chain rule (Stratonovich)

$$y_t = y_0 + \sum_i \int_0^t V_i \circ y_\tau dW_\tau^i$$

Itô lemma (stochastic chain rule) \Rightarrow

$$f \circ y_t = f \circ y_0 + \sum_j \int_0^t V_j \circ f \circ y_\tau dW_\tau^j$$

e.g.
$$V_i \circ y_t = V_i \circ y_0 + \sum_j \int_0^t V_j \circ V_i \circ y_\tau dW_\tau^j$$

$$y_t = y_0 + \sum_i \int_0^t \left(V_i \circ y_0 + \sum_j \int_0^{\tau_1} V_j \circ V_i \circ y_{\tau_2} dW_{\tau_2}^j \right) dW_{\tau_1}^i$$

Stochastic Taylor series

$$y_t = y_0 + \sum_i \int_0^t dW_{\tau_1}^i V_i \circ y_0 + \sum_{i,j} \int_0^t \int_0^{\tau_1} V_j \circ V_i \circ y_{\tau_2} dW_{\tau_2}^j dW_{\tau_1}^i$$

Now choose $f = V_j \circ V_i \Rightarrow$

$$y_t = y_0 + \sum_i \underbrace{\int_0^t dW_{\tau_1}^i}_{J_i(t)} V_i \circ y_0 + \sum_{i,j} \underbrace{\int_0^t \int_0^{\tau_1} dW_{\tau_2}^j dW_{\tau_1}^i}_{J_{ji}(t)} V_j \circ V_i \circ y_0 + \dots$$

- ▶ Feynman–Dyson path ordered exponential, Neumann series, Peano–Baker series, Chen–Fleiss series, stochastic B-series,...
- ▶ Euler-Maruyama and Milstein methods
- ▶ Need approximations for iterated integrals

Flow map/exponential Lie series

$$y_t = \varphi_t \circ y_0$$

$$\varphi_t = \text{id} + \sum_i J_i(t) V_i + \sum_{i,j} J_{ji}(t) V_j \circ V_i + \sum_{i,j,k} J_{kji}(t) V_k \circ V_j \circ V_i + \dots$$

Set $\varphi_t = \exp \psi_t$ then \Rightarrow

$$\begin{aligned} \psi_t &= (\varphi_t - \text{id}) - \frac{1}{2}(\varphi_t - \text{id})^2 + \frac{1}{3}(\varphi_t - \text{id})^3 + \dots \\ &= \sum_{i=0}^d J_i V_i + \sum_{i>j} \frac{1}{2}(J_{ij} - J_{ji})[V_i, V_j] + \dots \end{aligned}$$

- Magnus 1954, Chen 1957, Kunita 1980, Strichartz 1987, Ben Arous 1989, Castell 1993, Castell–Gaines 1995, Moan 2004

Castell–Gaines (ODE) method

Truncated exponential Lie series across $[t_n, t_{n+1}]$:

$$\hat{\psi}_{t_n, t_{n+1}} = \hat{J}_1 V_1 + \hat{J}_2 V_2 + \hat{J}_0 V_0 + \frac{1}{2}(\hat{J}_{12} - \hat{J}_{21})[V_1, V_2]$$

Approximate solution:

$$y_{t_{n+1}} \approx \exp(\hat{\psi}_{t_n, t_{n+1}}) \circ y_{t_n}.$$

Castell–Gaines: solve ODE

$$u'(\tau) = \hat{\psi}_{t_n, t_{n+1}} \circ u(\tau)$$

across $\tau \in [0, 1]$ with $u(0) = y_{t_n}$ gives $u(1) \approx y_{t_{n+1}}$.

Challenges?

- ▶ Combinatorial Hopf algebras
- ▶ Positivity preservation
- ▶ SPDEs and inference

Basic idea

$$\varphi_t = \text{id} + \sum_i J_i(t) V_i + \sum_{i,j} J_{ji}(t) V_j \circ V_i + \sum_{i,j,k} J_{kji}(t) V_k \circ V_j \circ V_i + \dots$$

$$\varphi = \sum_{w \in \mathbb{A}^*} J_w V_w$$

$$\begin{aligned} \psi &= \sum_{k=1}^{\infty} C_k (\varphi - \text{id})^k \\ &= \sum_{w \in \mathbb{A}^*} K_w V_w \end{aligned}$$

$$\mathbb{A}^* = \{\text{words over } \mathbb{A} = \{0, 1, \dots, d\}\}$$

Hopf algebra of words

Shuffle relations:

$$J_{a_1} J_{a_2} = J_{a_1 a_2} + J_{a_2 a_1}$$

$$J_{a_1} J_{a_2 a_3} = J_{a_1 a_2 a_3} + J_{a_2 a_1 a_3} + J_{a_2 a_3 a_1}$$

$$\begin{aligned} J_{a_1 a_2} J_{a_3 a_4} &= J_{a_1 a_2 a_3 a_4} + J_{a_1 a_3 a_2 a_4} + J_{a_1 a_3 a_4 a_2} \\ &\quad + J_{a_3 a_1 a_4 a_2} + J_{a_3 a_1 a_2 a_4} + J_{a_3 a_4 a_1 a_2} \end{aligned}$$

Shuffle product: $J_u J_v \longrightarrow u \sqcup v$

Two Hopf algebra structures, set:

$$\mathcal{H} = \mathbb{R}\langle \mathbb{A} \rangle \otimes \mathbb{R}\langle \mathbb{A} \rangle$$

$$(u \otimes x)(v \otimes y) = (u \sqcup v) \otimes (xy)$$

Pullback to Hopf algebra

$$\varphi = \sum_{w \in \mathbb{A}^*} w \otimes w$$

$$\begin{aligned} \Rightarrow \psi &= \sum_{k \geq 1} C_k (\varphi - 1 \otimes 1)^k \\ &= \sum_{k \geq 1} C_k \left(\sum_{w \in \mathbb{A}^+} w \otimes w \right)^k \\ &= \sum_{k \geq 1} C_k \sum_{u_1, \dots, u_k \in \mathbb{A}^+} (u_1 \sqcup \dots \sqcup u_k) \otimes (u_1 \dots u_k) \\ &= \sum_{w \in \mathbb{A}^*} \left(\sum_{k=1}^{|w|} C_k \sum_{\substack{u_1, \dots, u_k \in \mathbb{A}^+ \\ w = u_1 \dots u_k}} u_1 \sqcup \dots \sqcup u_k \right) \otimes w \end{aligned}$$

Sinh-log series?

- ▶ Diffusion vector fields commute: exponential Lie series has smaller mean-square error (orders $1/2$, 1 , $3/2$)
- ▶ Do not commute: counter-example for order 1
- ▶ *Assume no drift*: \Rightarrow sinh-log series:
 - ▶ $K_w \equiv \frac{1}{2}(J_w - J_{\alpha \circ w})$
 - ▶ Use concatenation-shuffle operator algebra
 - ▶ Remainder is smaller

Positivity preservation

Recall Heston model for pricing options:

$$\begin{aligned} du_t &= \mu u_t dt + \sqrt{v_t} u_t dW_t \\ dv_t &= \kappa(\theta - v_t) dt + \varepsilon \sqrt{v_t} dZ_t \end{aligned}$$

- ▶ Square-root diffusions prototypical Langevin dynamics
- ▶ Degrees of freedom $\nu := 4\kappa\theta/\varepsilon^2$
- ▶ Transition probability $P(v_t < x: v_0)$ is:

$$F_{\chi^2_\nu}(\lambda)(z) = \frac{e^{-\lambda/2}}{2^{\nu/2}} \sum_{j=0}^{\infty} \frac{(\lambda/2)^j}{j! 2^j \Gamma(\nu/2 + j)} \int_0^z \xi^{\nu/2+j-1} e^{-\xi/2} d\xi$$

where $z = x \hat{\eta}(t)$ and $\lambda = v_0 \eta(t)$ (see e.g. Andersen)

Zero boundary attracting, attainable?

i.e. solution of PDE $u_t = \mathcal{L}u$ is:

$$F_{\chi_\nu^2(\lambda)}(z) = \frac{e^{-\lambda/2}}{2^{\nu/2}} \sum_{j=0}^{\infty} \frac{(\lambda/2)^j}{j! 2^j \Gamma(\nu/2 + j)} \int_0^z \xi^{\nu/2+j-1} e^{-\xi/2} d\xi$$

Zero boundary, for:

- ▶ $\nu > 1$: non-attracting \Rightarrow implicit numerical methods*
- ▶ $\nu \leq 1$: attracting and attainable \Rightarrow transition prob

*with extreme caution depending on vector fields

Extension to SPDEs modelling biochemical reactions?

SPDEs

Gyöngy, Lord, Shardlow, Davie, Gaines, Stuart, Kloeden, . . .

- ▶ Brown report \Rightarrow major important challenge
- ▶ Stuart *et. al.* \Rightarrow path sampling, e.g. of bridge process
- ▶ Lots of physical/ocean/weather and defence applications

More challenges

- ▶ Numerical stability (Buckwar, ...)
- ▶ Symplectic methods (Tretyakov, Bou–Rabee, ...)
- ▶ Driving fractional Brownian motions (Baudoin, ...)
- ▶ Driving processes with jumps, eg. Lévy processes
- ▶ Backward stochastic differential equations (Protter, ...)

Introductory references

- ▶ Theory:

L.C. Evans: *An introduction to stochastic differential equations*

<http://math.berkeley.edu/~evans>

- ▶ Numerics:

D.J. Higham: *An algorithmic introduction to numerical simulation of stochastic differential equations*

SIAM Review 43 (2001), pp. 525–546

Itô or Stratonovich?

- ▶ Itô:

$$\int_0^T W_\tau^i dW_\tau^i = \frac{1}{2}(W_T^i)^2 - \frac{1}{2}T$$

- ▶ Stratonovich:

$$\int_0^T W_\tau^i \circ dW_\tau^i = \frac{1}{2}(W_T^i)^2$$

- ▶ Itô to Stratonovich:

$$V_0 = \tilde{V}_0 - \frac{1}{2} \sum_{i=1}^d V_i^2$$

- ▶ Stratonovich calculus familiar, easier, swapping to/back to Itô form trivial

Quadrature

How do we estimate $J_{12}(t_n, t_{n+1})$?

We can approximate it strongly using:

- ▶ Its conditional expectation $\hat{J}_{12}(t_n, t_{n+1})$
- ▶ Karhunen–Loeve expansion (Fourier expansion)
- ▶ Wiktorsson's method (most promising) — looks at the joint probability distribution function for J_1 , J_2 and J_{12}
- ▶ Stump and Hill's method (analogous)

Quadrature error I

With $\tau_q = t_n + q\Delta t$, $q = 0, \dots, Q - 1$, $Q\Delta t = h$

$$\begin{aligned} J_{12}(t_n, t_{n+1}) &= \int_{t_n}^{t_{n+1}} \int_{t_n}^{\tau} dW_{\tau_1}^1 dW_{\tau}^2 \\ &= \sum_{q=0}^{Q-1} \int_{\tau_q}^{\tau_{q+1}} W_{\tau}^1 - W_{t_n}^1 dW_{\tau}^2 \\ &= \sum_{q=0}^{Q-1} \int_{\tau_q}^{\tau_{q+1}} (W_{\tau}^1 - W_{\tau_q}^1) + (W_{\tau_q}^1 - W_{t_n}^1) dW_{\tau}^2 \\ &= \sum_{q=0}^{Q-1} J_{12}(\tau_q, \tau_{q+1}) + \sum_{q=0}^{Q-1} (W_{\tau_q}^1 - W_{t_n}^1) \Delta W^2(\tau_q) \end{aligned}$$

Quadrature error II

With $\tau_q = t_n + q\Delta t$, $q = 0, \dots, Q - 1$, $Q\Delta t = h$

$$\begin{aligned}\left\| J_{12}(t_n, t_{n+1}) - \hat{J}_{12}(t_n, t_{n+1}) \right\|_{L_2}^2 &= \sum_{q=0}^{Q-1} \left\| J_{12}(\tau_q, \tau_{q+1}) \right\|_{L_2}^2 \\ &= \sum_{q=0}^{Q-1} (\Delta t)^2 \\ &= Q(\Delta t)^2 \\ &= h^2/Q\end{aligned}$$

$$\Rightarrow \left\| J_{12}(t_n, t_{n+1}) - \hat{J}_{12}(t_n, t_{n+1}) \right\|_{L_2} = h/\sqrt{Q}$$

Wiktorsson: simulation of Lévy area I

- ▶ Conditional distribution of $\xi = J_{12}(h)$ given $J_1(h)$, $J_2(h)$:

$$\phi(\xi) = \frac{\frac{1}{2}h\xi}{\sinh(\frac{1}{2}h\xi)} \exp\left(-a^2\left(\frac{1}{2}h\xi \coth(\frac{1}{2}h\xi) - 1\right) + ih\xi b\right)$$

where $a^2 = (J_1^2(h) + J_2^2(h))/(2h)$ and $b = J_1(h)J_2(h)/2h$

- ▶ Need to sample from this

Wiktorsson: simulation of Lévy area II

- ▶ Karhunen–Loève $X_{ij}, Y_{ij} \sim N(0, 1)$:

$$A_{ij}(h) = \frac{h}{2\pi} \sum_{k=1}^{\infty} \frac{1}{k} \left(X_{ik} (Y_{jk} + \sqrt{\frac{2}{h}} J_j(h)) - X_{jk} (Y_{ik} + \sqrt{\frac{2}{h}} J_i(h)) \right)$$

- ▶ Truncate after Q terms, mean-square error $\mathcal{O}(h^2/Q)$
- ▶ Q large, tail sum \sim multivariate Gaussian distribution
- ▶ Approximate with Gaussian RV: mean-square error $\mathcal{O}(h^2/Q^2)$
- ▶ SDELab: Gilling & Shardlow

Stochastic integral properties

Stratonovich-to-Itô relations:

$$w = a_1 a_2: J_w = I_w + \frac{1}{2} I_0 \delta_{a_1=a_2 \neq 0}$$

$$w = a_1 a_2 a_3: J_w = I_w + \frac{1}{2} (I_{0a_3} \delta_{a_1=a_2 \neq 0} + I_{a_1 0} \delta_{a_2=a_3 \neq 0})$$

$$w = a_1 a_2 a_3 a_4: J_w = I_w + \frac{1}{4} I_{00} \delta_{a_1=a_2 \neq 0} \delta_{a_3=a_4 \neq 0} \\ + \frac{1}{2} (I_{0a_3a_4} \delta_{a_1=a_2 \neq 0} + I_{a_1 0a_4} \delta_{a_2=a_3 \neq 0} + I_{a_1 a_2 0} \delta_{a_3=a_4 \neq 0})$$

Expectations of Itô integrals: $\mathbb{E}(I_{\bullet j \bullet}) = 0$ for $j \neq 0 \Rightarrow$

$$\mathbb{E}(J_i) = 0$$

$$\mathbb{E}(J_{ij}) = \frac{1}{2} h \delta_{i=j \neq 0}$$

$$\mathbb{E}(J_{ijk}) = 0$$

$$\mathbb{E}(J_{ijkl}) = \frac{1}{8} h^2 \delta_{i=j \neq 0} \delta_{k=l \neq 0}$$

Basic idea

- ▶ Construct the new series $\psi_t = F(\varphi_t)$
- ▶ Truncate and use \hat{J}_i, \hat{J}_{ij} to produce $\hat{\psi}_t$
- ▶ Approximate flow-map is $\hat{\varphi}_t = F^{-1}(\hat{\psi}_t)$
- ▶ “Flow error” is the flow remainder $R_t = \varphi_t - \hat{\varphi}_t$
- ▶ Approximate solution is $\hat{y}_t = \varphi_t \circ y_0$
- ▶ Error/remainder is $R_t \circ y_0$
- ▶ Mean-square error measure is

$$\|R_t \circ y_0\|_{L^2}^2 \equiv \mathbb{E}(R_t \circ y_0)^T (R_t \circ y_0)$$